IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-8, Aug 2018 E-ISSN: 2347-2693

Psychological Stress Detection from Social Media Data using a Novel Hybrid Model

Shaikha Hajera[1], Mohammed Mahmood Ali[2*]

[1]MJCET,OU, Telangana -500034, INDIA.
[2]Osmania University, Telangana -500034, INDIA

**Corresponding Author: mahmoodedu@gmail.com.*

Available online at: www.ijcseonline.org

Accepted: 17/Aug /2018, Published: 31/Aug/2018

Abstract: Psychological stress is a biggest threat to human's health. Hence, it is vital to detect and manage stress before it turns into severe problem. However, conventional stress detection strategies rely on psychological scales and physiological devices, which require active individual participation making it labor-consuming and expensive. With the rapid evolution of social media networks, people are willing to sharetheir everyday events and moods via social media platforms, making it practicable to leverage this online social media content for stress detection as these data timely reflect user's real-life emotional state. To automatically predict stress, we have defined a set of stress-related textual 'F = {f1, f2, f3, f4}', visual 'vF = {vf1, vf2}', and social 'sf' features, and thenproposed a hybrid model Psychological Stress Detection (PSD) - a Probabilistic Naïve Bayes Classifier combined with Visual (Hue, Saturation, Value) and Social modules,to leverage text, image and social interaction information for stress detection from social media contentExperimental results show that the proposed PSD model improves the detection performance,when compared to TensiStrength and Teenchat frameworkPSD achieves 95% of Precision rate. PSD model would be useful in developing stress detection tools for mental health agencies and individuals.

Keywords: Psychological Stress Detection; Social Media interaction; Health agencies; Physiological Signals*;*

I. INTRODUCTION

Psychological stress is considered the biggest threat to an individual's health. Three of every hundred people in metropolitan areas are assessed to suffer from stress. Long-lasting stress may lead to many severe physical and mental problems, such as depressions, insomnia, and even suicide. Stress and suicide are closely interlinked. At its worst, anxiety can lead to suicide [1]. According to World Health Organization (WHO) over 56 million Indian's suffer from depression and Corporate sector in India as well reported an increase in stress over the last two years, a survey by workplace solutions provider Regus in 2015, reported 57% of corporate India is under stress and about one in five people in the country need counseling, either psychological or psychiatric. The National Institute of Mental Health and Neurosciences (NIMHANS) published a mental health assessment said 5% of the population suffers from depression as of 2016. Thus, recognizing stress or depression at an early stage is critical for reducing suicidal deaths and deliberate self-harm across the spectrum. A range of sensors and non-textual methods have been developed to detect stress as sensors can offer data regarding the intensity and quality of a person's internal affect Experience [7]. For examples, H. Kurniawan presents stress detection from speech and galvanic skin response signals [3], and [6].S. Greene

Fig 1: Sample Post containing text, image and social interactions (i.e., comments and likes)

describes the stress detection method utilizing sensor which monitors Heart Rate Variability (HRV) [3], [5], and [7]. These sensors are expensive and most of the Traditional stress detection methods are mainly based on one-on-one interviews conducted by psychologists or self-report questionnaires [13] or wearable sensors[16]. However, these are usually labor & time-consuming and are reactive methods. The query is, are there any appropriate and practical methods for stress recognition?

The rapid growth of Social Media is Changing People's Life, as Well as Research in Healthcare and Wellness. Report from Global Social Media research 2018[4] shows there are in total 3.196 billion active social media users worldwide and

© 2018, IJCSE All Rights Reserved

among them Facebook is the most popular social network with total 2.01 billion active users monthly and Twitter is the fastest growing social networks with total of 328 million active users. With the quick advancement of social networking sites, individuals are excited to use it as a platform to express their moods and day-to-day life actions making it a popular platform to express thoughts via posts. People share thoughts, express emotions, record daily habits via textual and image posts and also interconnect with friends, a sample is depicted in Figure 1. As these social media data reveal user's reallife situations and emotions in a timely fashion, it offers new prospects for assessing, modeling, and mining user's activity patterns through social networks, such social data can find its theoretical basis in psychology research and Encoding emotional information in text is a common practice especially in online interactions [2]. Thus, we can obtain linguistic and visual content from individual's posts that may indicate stress related symptoms that makes the detection of user's psychological stress feasible through their posts, posting behavior and social interaction on micro-blog or social media.

The rest of paper is organized as follows: The Section II gives an overview of the related work and also describes the problem statement. Section III introduces the definitions of the proposed features and presents the hybrid model Psychological Stress Detection (PSD) for stress recognition and describes the algorithmic structure and modules of the proposed PSD model. Section IV presents experimental results. Lastly, Section V provides the conclusion and discusses future scope.

II. Related Work And Problem Defination

There have been research efforts on harnessing social media data for developing mental and physical healthcare tools. K. Lee et.al's [8] proposed to leverage social media data for real-time disease investigation; while [10] tried to link the vocabulary gaps between health seekers and providers using the community generated health data. Psychological stress detection is interrelated to the topics of sentiment assessment and emotion recognition. Our recent work [25], [35] presented relative analysis of various psychological stress recognition Approaches. Study on Emotion Detection in Social Networks, Computer-aided recognition, scrutiny, and application of emotion, especially in social media, have drawn considerable attention in current years [11], [12], and [15]. Associations between psychological stress and personality traits can be a thought-provoking issue to contemplate [17], [18], and [20]. Several studies on social networks based emotion analysis uses text-based linguistic features and typical classification approaches. As discussed above a range of sensors and non-textual methods have been developed to detect stress but they are expensive. With the development of social networks people are willing to share their daily events and moods, and interact with friends through the social networks, making it possible to leverage online social network data for stress detection.

There are also some exploration [22], and [27] using user posting contents on social network to identify user's stress revealed that leveraging this social media data for healthcare, and in particular stress detection, is feasible. However, these mechanisms mainly leverage the textual contents and consider only single post of an individual in social networks. In certainty, data in social networks are typically composed of chronological and inter-connected items from diverse sources i.e., it also contains images posts apart from textual posts, making it be actually multi-media data and single post reflects the instant emotion but people's emotion or psychological stress states are habitually more persistent, fluctuating over different time periods. In recent years, extensive study started to focus on emotion detection in social networks from sequential post series [21], [23], and [24]. Our work is to leverage sequential post (multi-media data) content over specific sampling period to detect stress. The proposed novel hybrid model Psychological Stress Detection (PSD) defines set of stress-related textual 'F = {f1, f2, f3, f4}', visual 'vF = {vf1, vf2}', and social 'sf' features that makes the detection efficient and effective leveraging textual and visual post content.

Before presenting our problem definition, let's first define necessary notations. Let P be a set of users posts on a social media network i.e., pi ∈ P. Let U be set of users. Each user ui∈ U, posts on social media containing text, image. Where i denotes number of posts and users that can range from 1 to 'n' i.e., i = {1, 2, 3, …, n}.

- ***Stress state*:** stress state denoted by's' of user ui∈ U at time t is represented as a triple (s, ui, t), or $\mathbf{s_i^t}$. In the study, a binary stress state $\mathbf{s_i^t}$∈ {0, 1} is considered, where $\mathbf{s_i^t}$ = 1 specifies that user ui is stressed at time t, and $\mathbf{s_i^t}$ = 0 specifies that the user is non-stressed at time t.
- ***Stress feature set:*** Let F, vF and sf be the stress feature sets in study i.e., F = {f1, f2, f3, f4}, vF ={vf1,vf2}, sf , where f1 – f4 are linguistic stress lexicon features, vf1 and vf2 are visual features and sf is social feature as shown above in table 4.2.(page number 19).
- ***Problem Statement:***To Detect Psychological Stress state 's' Given a series of user posts P. To address this problem we first define the stress feature sets they are Textual feature (F), Visual feature (vF) and social feature (sf) then propose Psychological Stress Detection (PSD) model to leverage the social media content for stress detection.

© 2018, IJCSE All Rights Reserved

III. Proposed Psychological Stress Detection (PSD) Model

A. Feature Categorizations

To cope with the problem of psychological stress detection we have defined set of stress-related textual 'F = {f1, f2, f3, f4}', visual 'vF = {vf1, vf2}', social 'sf' features. Table 1 shows categorization of these features. Textual Feature is a set of four features, F = {f1, f2, f3, f4}, these feature stores word count of a particular sentence comparing them with linguistic stress lexicon explained in Table 1. 'f1' stores number of stress and negative words present in the sentence posted. 'f2' stores number of positive emotion words present in the sentence posted, 'f3' stores negative emoticons if present and 'f4' stores negating words present in the sentence posted. These features help the probabilistic module in calculating the probability of each class and labeling the textual post to its particular class as shown in Figure 4.

Visual Feature is a set of two features, vF = {vf1, vf2} they stores the mean value of the brightness and saturation, which is calculated in the visual module algorithm as shown in Figure 5. Value of Image brightness and saturation will be compared with threshold value to categorize the particular image post as stress or non-stress as shown in Figure 3. Social Feature 'sf' stores the mean value of number of likes a particular post gets it give us the social attention degree that means when user posts are negative or stressful it gets more attention from the friends of that particular user via likes or comments. Thus, we will find the mean of number of likes and categories the posts accordingly as shown in Figure 3.
The detail Architecture of the model is shown in the Figure 2. Psychological stress detection has many challenges to overcome them first have categorized features as discussed in section III.B, now will design modules to utilize these features for stress detection. Probabilistic module fetches textual feature 'F = {f1, f2, f3, f4}', from content of user's post comparing with Linguistic Stress Lexicon. Visual module fetches visual features 'vF = {vf1, vf2}' from image post, and Social module will fetch social attention feature 'sf' as discussed in this section .

B. Architecture

Figure 2 shows the architecture of our model. There are three types of information that we can use as the initial inputs, i.e., Textual features, Visual features and Social features whose detailed computation will be described later. We address the solution through the following three key modules:

1. Design the Probabilistic module to fetch textual feature 'F = {f1, f2, f3, f4}', from content of user's post comparing with Linguistic Stress Lexicon.
2. Visual module fetches visual features 'vF = {vf1, vf2}' from image post, and
3. Social module fetches social attention feature 'sf' as discussed in section III.B.c.

Steps involved in the architecture of the proposed system are as follows:

i. The first step captures the Post and messages sent between the users of Social Networking Site (SNS). These messages are saved in the database.
ii. In this step, a text post is taken and this text message is converted to plain text removing stop-words (such as articles, preposition). Next, will check if the user has also posted an image if yes, then will send it to visual module to process its mean value of brightness and saturation. Next, will check social attention feature to calculate the mean of likes.
iii. This is the main step in which will compare the plain text with linguistic stress lexicon to find out the number of stress words, negative words present in the text and store the result into feature set denoted by 'f1-f4'. Next, will extract the image features and store the result into feature set denoted by 'vf1' and 'vf2'. Lastly will calculate the number of likes and stored it in feature 'sf' as shown in Table 1.
iv. The Textual (linguistic) features are fed into the probabilistic module to predict the class of the text post. The module uses maximum likelihood estimates for the detection. The visual features are fed into the visual module, where the brightness and saturation of the image is calculated as shown in figure 5 and compared with the threshold value as shown Figure 3. Social feature is fed into social module to mean of likes and to compare it with a threshold value shown in Figure 3.
v. In the last step, results are displayed to the user if the user posts satisfies all the threshold values then only his/her posts will be tagged as stress as shown in Fig. 3 and if the user is status is stressed then will send the recommendation to the user regarding stress management.

a) Probabilistic Module

To fetch textual feature from the posts will make use of Naïve Bayes classifier which is based on Bayes theorem [31]. This classifier algorithm uses conditional independence, means it assumes that an attribute value on a given class is independent of the values of other attributes. It is simple yet efficient classifier that is used to classify uncertain data. It works on the principles of Bayes Probability theorem that means it will be having a set of known results for some words and it compares those results with the words in a particular sentence.The Bayes theorem is as follows: Let X = $\{x_1, x_2, \dots, x_n\}$ be a set of 'n' attributes. In Bayesian, 'X' is considered as evidence and 'H' is some hypothesis means, the data of X belongs to specific class 'C'. We have to

© 2018, IJCSE All Rights Reserved

determine $P(H/X)$, the probability that the hypothesis H holds given evidence i.e. data sample X [30]. According to Bayes theorem the $P(H/X)$ is expressed as shown in equation 1 or 2.

$$P\left(\frac{H}{X}\right) = P(X/H)P(H)/P(X) \qquad (1)$$

or

$$\mathrm{P(Sentence/Class_i)} = \prod_{i}^{n} \mathrm{P(word_n/Class_i)} \qquad (2)$$

Where $Stress_i$ is any one of the class in the list i.e., 'No-Stress', 'Stress' and Neutral. $Word_m$ - are Number of word in the given sentence, here wordn = fn, where n = 4 refer Table1

Table 1: Illustrates set of knowledge based pre-defined logical rules.

Short name	Sample Words/ Images to be detected	Feature notation	Description
Textual feature (F): Rule 1 – Algorithm 1 (Figure 3) will check threshold values to tag the textual posts as stress			
Stress Lexicon	stress,depress,anxious,sad,depress,anger,pain,restless,sleepless,failure,sorrow,lonely,quit,worry,nervous,mental,break,anguish,argu,arrogant,atrocious,bereave,brokenhearted,Hurt, ugly, anger, annoy, boredom, sadness, grief, disapproval, low self-esteem, callous, despair, devastate, hate……..	f1	Number of stress words and negative words.
Positive emotion Lexicon	Love, joy, trust, surprise, care, happy, amazement, interest…..	f2	Number of positive words.
Negative Emoticons	"%-(", ")-:", "):", ")o:" , ":(", ":*(", ":,(", ":-", ":-&", ":-("…….	f3	Number of popular Negative emoticons.
Negating words Lexicon	Cannot, cant, couldn't, didn't, doesn't, don't, hasn't, isn't, never, no, not, shouldn't, won't, wouldn't.	f4	Number of negating words.
Visual feature (vF): Rule 2 - Algorithm 1 (Figure 3) will check threshold values to tag Visual posts as stress			
Brightness		vf1	Mean of image pixels i.e., brightness
Saturation		vf2	Mean of saturation.
Social feature(sf): Rule 3 - Algorithm 1 (Figure 3) will check threshold values			
Social attention	Likes (9)	sf	Number of likes and mean value of it.

If textual feature fetched from the given sentence matches the words in the linguistic stress lexicon explained in Table 1 will assign class to that sentence according to the probability score of each class. The detail computation of probabilistic module is shown in Probabilistic Module Algorithm Figure 4. As discussed in section III.B the textual Feature are a set of four features, F = {f1, f2, f3, f4}, these feature stores word count of a particular sentence comparing them with linguistic stress lexicon explained in Table 1. 'f1' stores number of stress and negative words present in the sentence posted. 'f2' stores number of positive emotion words present in the sentence posted, 'f3' stores negative emoticons if present and 'f4' stores negating words present in the sentence posted. These features help the probability module in labeling the posted sentence is having stress or non-stress by calculating the probability of each class.

© 2018, IJCSE All Rights Reserved

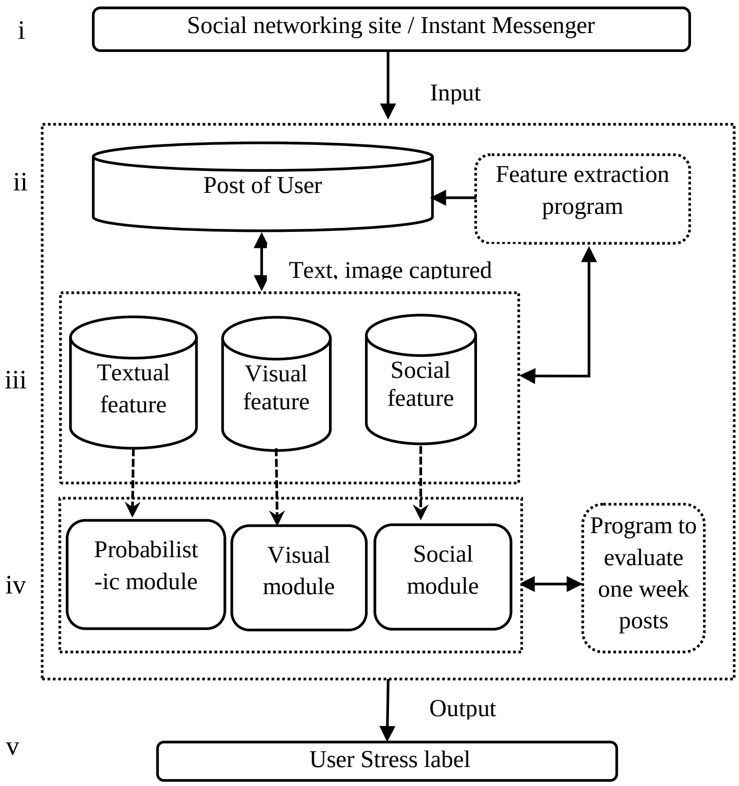

Fig.2: Psychological Stress Detection (PSD) Architecture.

b) VISUAL MODULE

Based on previous work on color psychology theories [28], we combine the following features as the visual middle-level representation: - Saturation: the mean value of saturation and its contrast. It describes the colorfulness and the differences of an image. Psychological experiments in S. R. Ireland's [29] find out that people under stress and anxiety prefer lower saturation than normal states, revealing the correlation between stress and saturation of images. Brightness: the mean value of brightness and its contrast. It illustrates the perception elicited by the luminance and the related differences of an image (e.g., low brightness makes people feel negative, while high brightness elicits mainly positive emotional associations). To calculate image brightness first, we need to (briefly) analyze what is the result of the average value of the sum of the RGB channels. For humans, it is meaningless. Is pink brighter than green ? I.e., why would you want (0, 255, 0) to give a lower brightness value than (255, 0, 255) ? Also, is a mid gray (128, 128, 128) bright just like a mid green (128, 255, 0) ? To take this into consideration, I only deal with the color brightness of the channel as is done in the HSV color space [32], [14]. This is simply the maximum value of a given RGB triplet. The rest is heuristics. Let max_rgb = max (RGB-i) for some point i. If max_rgb is lower than 128 (assuming a 8bpp image), then we found a new point i that is dark, otherwise it is light. Doing this for every point i, we get A points that are light and B points that are dark. If (A - B)/(A + B) >= 0.5 then we say the image is light. Note that if every point is dark, then you get a value of -1, conversely if every point is light you get +1. The previous formula can be tweaked so you can accept images barely dark. In the code I named the variable as fuzzy, but it does no justice to the fuzzy field in Image Processing. So, we say the image is light if (A - B)/(A + B) + fuzzy >= 0.5.Following is the mathematical relationship between RGB space to HSI (hue, saturation, and intensity) or Hue, Saturation, Value (HSV) color space model[33], [34].

- Brightness or Intensity formula

$$vf1 = \frac{(R+G+B)}{3} \tag{3}$$

- Saturation formula

$$vf2 = 1 - \frac{3}{(R+G+B)}(\min(R,G,B)) \tag{4}$$

Where vf1 is brightness or Intensity and vf2 is saturation as shown in equation 3 and 4. The detail computation of visual module is shown in Visual Module Algorithm figure 5. Visual Feature is a set of two features, vF = {vf1, vf2} they stores the mean value of the brightness and saturation, which is calculated in the visual module algorithm. Value of Image brightness and saturation will be compared with threshold value to categorize the particular image post as stress or non-stress as shown in Figure 3.

c) SOCIAL MODULE

Besides the text content and image content of a post, some additional features such as likes can also imply one's stress state to some degree. We define a post's social attention degree based on these additional features into social attributes. As apparently stressful posts may attract more attention from friends. The number of comments, likes reveals how much attention a post attracts. To find out the changes of attention degree of one's post, we first calculate the sample mean sf shown in equation 5, all M - values M_i and total number of items in the samples.

$$Mean\overline{sf} = \frac{(\sum M_i)}{n} \tag{5}$$

Social Feature 'sf' stores the mean value of number of likes a particular post gets it give us the social attention degree that means when user posts are negative or stressful it gets more attention from the friends of that particular user via likes or comments. Thus, we will find the mean of number of likes and categories the posts accordingly as shown in Figure 3.

Algorithm 1: *Psychological stress Detector (PSD) Algorithm*
Require: $\boldsymbol{p_i}$, $\boldsymbol{u_i}$, *(Social media posts **P** of users **U**).*

© 2018, IJCSE All Rights Reserved

Ensure: 's'*(Detection ofStress state i.e., stress or non-stress)*

```
Initialize P, U, s = {stress , non-stress}, F ,vF ,sf .
u = getusername() //getusername function returns
                         current user name
While(i < U and P != NULL ) do
   F = Stress (P) //stress function of probability module
   calculates stress from     textualposts and return
   stress state
   vF = getImageLightness (image)
   //getImageLightness function of visual module
   calculates and return image brightness
   sf = num_likes() //num_likes function of social
   module calculates and return number  of likes and
   mean value of the likes.
   if(F = 1 and  vF<= 0.5 or sf>= 5) then //To display
   stress status for a week.
      stress_status =  stress  //tag that user having
                                       stress
         //if the user have more than five posts tagged as
         stress in a week   then system will display result
         as stress.
          if(loggedin_user == u and stress_status>= 5 )
          then
               s = stress //stress state
          else
               s = non-stress //stress state
          end if
   end if
end while
```

Fig.3: Psychological Stress Detection (PSD) Algorithm.

Algorithm 2: *Probabilistic Module Algorithm*
Require:*sentence* *(Textual post passed by PSD algorithm)*
Ensure: *stress_class**(textual post belongs to which class i.e., stress(x), non-stress(y) and neutral(z)).*

```
Function Stress (sentence )
   //given training data D which consists of database of
   words belonging to different class i.e., stress (x), non-
   stress (y) and neutral(z)
Initialize stress_class = {x,y}
//Implication that the analyzed text has 1/3 chance of
being in either of the 3 categories
 x = 0.333
 y =0.333
 z =0.334
 //find total number of word frequency of each class
 n = total number of word frequency
 //find conditional probability of words occurrence
 given a class
 P(fm / x) = wordcount / n(x)
 P(fm/y) = wordcount / n(y)
 P(fm/z) = wordcount/n(z)
 //classify a new sentence based on probability
 P(sentence / stress_class) = product-of[p(x) * P(fm/
 stress-class)]
 //assign sentence to class that has higher probability
 If ( stress_class == x )then
     return 1
 else
     return 0
 end if
End
```

Fig.4: Probabilistic Module Algorithm.

Algorithm 3:*Visual Module Algorithm*
Require: *'y' (visual post passed by PSD algorithm)*
Ensure: *'vF' (brightness and saturation value)*

```
Function getImageLightness (y)
 Initialize imageData =
 y.getImageData(0,0,width,height)
 Initialize data = imageData
 Initialize r,g,b,vf1,vf2,avg
 for(x=0 , len=data.length; x <len ; x+=4)do
      r = data[x]
      g = data[x+1]
      b = data[x+2]
```

© 2018, IJCSE All Rights Reserved

```
        vf1 = (r+g+b)/3
        vf2 = 1-[(3/(r+g+b))*(min(r,g,b))]
        avg += vf1+vf2
    end for
    return avg/255
nd
```

Fig.5: Visual Module Algorithm.

IV. Experimental Results

The proposed PSD model is compared with Mike Thelwall'sTensiStrength [27] which uses a lexical approach with lists of terms related to stress and non-stress. TensiStrength's terms are not only synonyms for stress, anxiety and frustration but also terms related to anger and negative emotions because stress can be a response to negative events and can cause negative emotions. It also attempts to detect the opposite state to stress, non-stress, through a parallel approach. This method is mainly based on text data in the social media, whereas other correspondingly significant content, like images and social actions are ignored and it is capable of detecting stress only from a single textual post, while single post reflects the instant emotion expressed, people's emotion or psychological stress states are usually more persistent and changes over different time periods and social media allows users to express their mental thoughts or feelings through not only textual posts but also in visual post i.e., through images. The proposed PSD is designed as such it takes into consideration textual and multi-media (images) data as well to detect stress. The sample scenario assumes at-most two posts (texts plus emoji plus image) per day and results are analyzed we achieved 83.3% of accuracy and 88.8% of F1-score. To further analyze the model we have taken more scenarios. In the Scenario2 we have considered post including emojis i.e., Text + Emoji and analyzed posts for a week and achieved 91.6% of accuracy and 92.2% of F1-score. In scenario3 we have considered post which includes emojis and image along with text post i.e., Text + Emoji + image shown in Table 2 and based on these these parameters we have achieved highest F1-score of 94% and accuracy of 91.9%. For further test our model, we have considered data from our application's database which is real time content and Table 3 shows the details of that data. The collected data consists of 1,920 posts based on these data we have evaluated PSD model and it showed an increased accuracy compared to J. Huang's Teenchat [22] and M.thelwall'sTensiStrength[27] system for stress detection. The figure 7 shows a comparison of efficiency of the PSD model from the result shown in table 4 we see that PSD model achieved better detection performance and showed F1-Score of 95.7% whereas TensiStrength shows 5.9% of less F1-score i.e., 90.10% and J. Huang's Teenchat [22] shows 15.5% of less F1-score i.e., 80.20% and when compared to our PSD model. The Table 4 shows the comparison of results using different parameters, as shown Tensistrength [27] achieved 89% of accuracy and it supports only text parameter for stress detection and for the same. J. Huang's Teenchat [22] model achieved 91.5% of accuracy based on all three parameters but the time taken by CPU is much higher than our PSD model which decreases its efficiency, our PSD model achieved 93% of accuracy when all parameters are considered and also take very less CPU time to process and detect stress from social media posts of individual users.

Table 2: Details of Data taken for result analysis.

Details	Count
Number of posts	1920
Number of Users	20
Number of weeks	8
Posts per week	12
Precision	95%

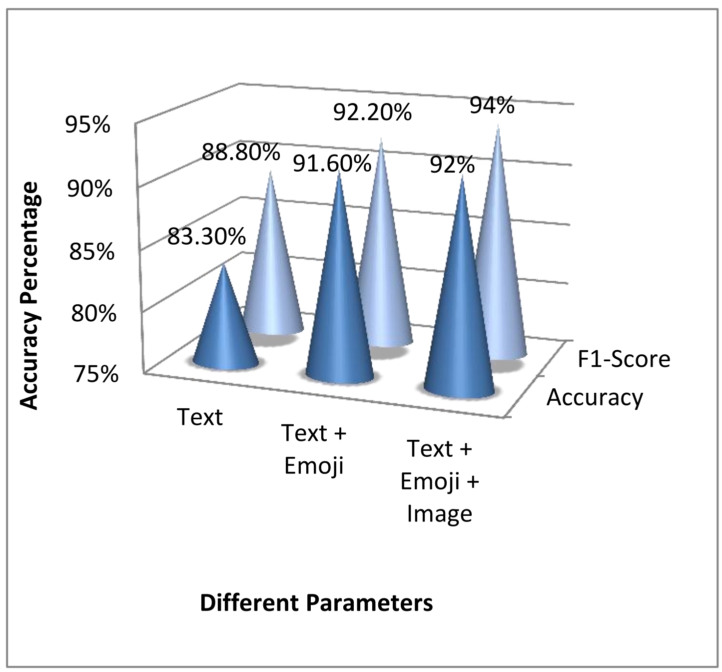

Fig.6: Comparison of F1-score and Accuracy based on different parameters.

Table 3: Shows a scenario of user's posts and result analysis (Text + Emoji +Image).

S.no.	Posts	Bayes Score	Image	I-score	Class
1	My life is *finished* very *alone* without my parentss*ad* feeling, in *pain*.	0.727		0.252	S
2	I am in *pain*, can't take it anymore *hate* life *depressed*.	0.8		0.237	S
3	Its*sad* when someone you know becomes you knew *depressed* to *lose* my friend	0.7		0.116	S
4	I *feeltense*, *anxious* or have *nervous* indigestion.	0.72		0.419	S
5	I can't do this anymore. My body is physically *incapable* to take this *pain*	0.67		0.165	S
6	I am so *disturbed*. Life is *hard*. *:(, %-(*	0.666		0.419	S
7	This *pain* is never ending *hate* this feeling *sad*, *dishearten*.	0.728		0.376	S
8	I am separated from my husband, i am *alonesad* and *depress*.	0.67		0.331	S
9	My sister came from USA happy to see her.	0.25		0.651	NS
10	I got a best singer award happy to receive this honor	0.1		0.7	NS
11	I lost my job dont have money feels demotivated cant take this mental pressure.	0.444		0.5	NS

Analysis	
Total posts to be detected as stressed	**9**
Number of posts detected as stressed	8
Precision	**99.9%**
Recall	**88.8%**
Accuracy	**91.9%**
F1-score	**94%**

Table 4: Comparison of efficiency & effectiveness using different model

Parameter / Models	Text	Text + Emoji	Text + Emoji + Image	Accuracy	F1-Score	CPU time
Teenchat[28]	✓	×	×	0.784	0.802	28s
TensiStrength [39]	✓	×	×	0.891	0.901	180s
PSD model	✓	✓	✓	0.93	0.957	35s

V. Conclusion And Future Scope

Psychological Stress is an increasing problem in the world, it is essential to monitor and control it regularly. In this paper, we presented a Model for detecting user's stress status from user's daily social media data, leveraging post's content along with user's social interactions. The inefficiency of the conventional methods [5], [16], [19], [22], and [27] motivates to adopt a novel technique which is transparent and inexpensive. Thus, the proposed strategy utilizes social media data and social interaction content to detect psychological stress employing a Novel Hybrid model Psychological Stress Detection (PSD). To completely leverage user's posts content, we proposed a hybrid PSD model which combines the probabilistic module (Bayes classifier) with the visual module (HSV model). In our model the user's stress states are monitored daily, weekly, bi-weekly or monthly. Hence, improving

© 2018, IJCSE All Rights Reserved

the accuracy of stress detection. Experimental results show that the proposed model can improve the detection performance when

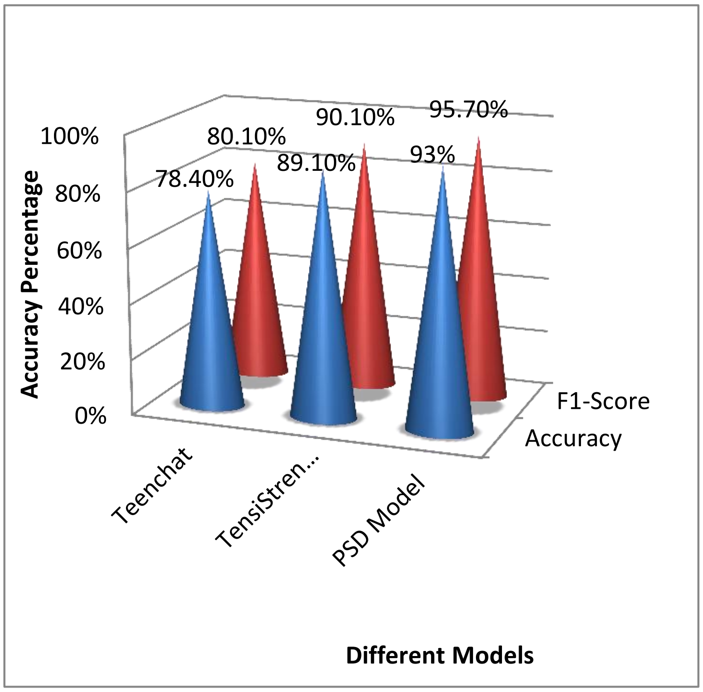

Fig.7: Comparison of efficiency of the models based on F1-score and Accuracy

compared to J. Huang's Teenchat [22], and TensiStrength [27], that achieved 95% of precision and 93% of accuracy.

Recent studies show apart from leveraging cross-media content and social interactions, Social Connectivity also plays a vital role in detecting stress states of an individual, which is a useful reference for future related studies. Some more useful references for future related studies:Need to explore our research on stressor responsible for stress as it is noteworthy to know the type of stress the user is dealing. Support for short form and slang language used in text post to be included and Support for Multilingual languages to be included.

REFERENCES

[1] J. Herbert, "Fortnightly review: Stress, the brain, and mental illness", British Medical J., pp. 530–535, Vol. 315, No. 7107, 1997.

[2] Liew, and Jasy Suet Yan, "fine-grained emotion detection in microblog text", Dissertations – ALL. pp 440, 2016.

[3] F.-T. Sun, C. Kuo, H.-T. Cheng, S. Buthpitiya, P. Collins, and M. L. Griss, "Activity-Aware Mental Stress Detection Using Physiological Sensors", In Proc. Of Intl. Conf. on Mobile Computing, Application, and Services (MobiCASE), Santa Clara, CA, 2010.

[4] Chaffey, D. (2016). "Global social media research summary 2016." Retrived in June 15th 2016 from http://www.smartinsights.com/social-mediamarketing/social-media-strategy/new-global-social-mediaresearch.

[5] H. Kurniawan, A.V. Maslov, and M. Pechenizkiy, "Stress detection from speech and galvanic skin response signals", in: Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems, pp. 209–214, 2013.

[6] V.P. Patil, K.K. Nayak, and M. Saxena, "Voice stress detection", Int. J. Electr. Electron. Comput. Eng., pp. 148–154, 2013.

[7] S. Greene, H. Thapliyal, and A. C. Holt, "A survey of affective computing for stress detection: Evaluating technologies in stress detection for better health", IEEE Consum. Electron. Mag., Vol. 5, No. 4, pp. 44–56, 2016.

[8] K. Lee, A. Agrawal, and A. Choudhary, "Real-time disease surveillance using twitter data: Demonstration on FLU and cancer", in Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, pp. 1474–1477, 2013.

[9] Sztajzel, "Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system", Swiss Med. Weekly, pp. 514–522, 2004.

[10] L. Nie, Y.-L. Zhao, M. Akbari, J. Shen, and T.-S. Chua, "Bridging the vocabulary gap between health seekers and healthcare knowledge", IEEE Trans. Knowl. Data Eng., vol. 27, no. 2, pp. 396–409, Feb. 2015.

[11] G. Coppersmith, C. Harman, and M. Dredze, "Measuring post traumatic stress disorder in twitter", in Proc. Int. Conf. Weblogs Soc. Media, 2014, pp. 579–582.

[12] R. Fan, J. Zhao, Y. Chen, and K. Xu, "Anger is more influential than joy: Sentiment correlation in weibo", PLoS One, vol. 9, 2014, Art. no. e110184.

[13] Pincus T, Swearingen C, Wolfe F "Toward a multidimensional Health Assessment Questionnaire (MDHAQ): assessment of advanced activities of daily living and psychological status in the patient-friendly health assessment questionnaire format", Arthritis Rheum 1999; 42: 2220-30

[14] W. Chen, Y. Q. Shi, and G. Xuan, "Identifying computer grahics using HSV color model and statistical moments of characteristic functions", In Proceedings of IEEE International Conference on Multimedia and Expo, Jul. 2007, pp. 1123–1126.

[15] Y. Zhang, J. Tang, J. Sun, Y. Chen, and J. Rao, "Moodcast: Emotion prediction via dynamic continuous factor graph model", Proc. IEEE 13th Int. Conf. Data Mining, 2010, pp. 1193–1198.

[16] A. Sano, and R. W. Picard, "Stress recognition using wearable sensors and mobile phones", In Proceedings of ACII, pp. 671–676, 2013.

[17] G. Farnadi, et al., "Computational personality recognition in social media", UserModel User-Adapted Interaction, vol. 26, pp. 109–142, 2016.

[18] J. Golbeck, C. Robles, M. Edmondson, and K. Turner, "Predicting personality from Twitter", in Proc. IEEE 3rd Int. Conf. Privacy, Security, Risk Trust, IEEE 3rd Int. Conf. Soc. Comput., 2011, pp. 149–156.

[19] A. Fernandes, R. Helawar, R. Lokesh, T. Tari, and A. V. Shahapurkar, "Determination of stress using blood pressure and galvanic skin response", In Proceedings of the International Conference on Communication and Network Technologies (ICCNT'14), pp. 165–168, Sivakasi, India, 2014.

[20] B. Verhoeven, W. Daelemans, and B. Plank, "Twisty: A multilingual twitter stylometry corpus for gender and personality profiling", in Proc. 10th Int. Conf. Language Resources Eval., PP. 1632–1637 2016.

[21] F. A. Pozzi, D. Maccagnola, E. Fersini, and E. Messina, "Enhance user-level sentiment analysis on microblogs with approval relations", in Proc. 13th Int. Conf. AI* IA: Advances Artif. Intell., PP. 133–144, 2013.

[22] J. Huang, Q. Li, Y. Xue, T. Cheng, S. Xu, J. Jia, and L. Feng.: "Teenchat: a chatterbot system for sensing and releasing adolescents' stress", In: X. Yin, K. Ho, D. Zeng, U. Aickelin, R.

© 2018, IJCSE All Rights Reserved

Zhou, H. Wang. (eds.) HIS LNCS, Vol. 9085, pp. 133–145. Springer, Heidelberg, 2015.
[23] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, and P. Li, "User-level sentiment analysis incorporating social networks," in Proc. SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2011, pp. 1397–1405.
[24] J.W. Pennebaker, R.J. Booth, and M.E. Francis, "Linguistic Inquiry and Word Count (LIWC)", 2007.
[25] Shaikha Hajera and Mohammed Mahmood Ali, "A Comparative Analysis of Psychological Stress Detection Methods", In IJCEM, Vol. 21 Issue 2, March 2018.
[26] Y.R. Tausczik, and J. W. Pennebaker, "The psychological meaning of words: LIWC and computerized text analysis methods", Journal of Language and Social Psychology, pp. 24-54, 2010.
[27] Thelwall M. "TensiStrength: stress and relaxation magnitude detection for social media texts", J Inf Process Managpp: 106–121, 2017.
[28] X. Wang, J. Jia, H. Liao, and L. Cai, "Affective image colorization", Journal of Computer Science andTechnology, Vol. 27, No. 6, pp. 1119-1128, 2012.
[29] S. R. Ireland, Y. M. Warren, and L. G. Herringer, "Anxiety and color saturation preference", Perceptualand Motor Skills, Vol. 75, pp. 545-546, 1992.
[30] S.-B. Kim, K.-S. Han, H.-C. Rim, and S. H. Myaeng. Some effective techniques for naive bayes text classification. IEEE Transactions on Knowledge and Data Engineering, 18(11):1457–1466, Nov. 2006.
[31] Z. Pawlak, "Rough sets, decision algorithms and Bayes's theorem," Eur. J. Oper. Res., Vol. 136, , Issue.1, pp. 181–189, 2002.
[32] C.P.Patidar, Meena Sharma, VarshaSharda, "Detection of Cross Browser Inconsistency by Comparing Extracted Attributes", International Journal of Scientific Research in Computer Science and Engineering, Vol.5, Issue.1, pp.1-6, 2017.
[33] MM Ali, MAU Rahman, S Hajera "A comparative study of various image dehazing techniques" International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), IEEE, pp. 3622-3628, 2017
[34] G. Sharma, A. Kumar, "REVIN: Reduced Energy Virtuous Immune Network for WSN", International Journal of Scientific Research in Computer Science and Engineering, Vol.5, Issue.2, pp.1-8, 2017.
[35] MM Ali, M Tajuddin, M Kabeer "SDF: psychological Stress Detection Framework from Microblogs using Pre-defined rules and Ontologies", International Journal of Intelligent Systems and Applications in Engineering, Vol.6, Issue.2, pp. 158-164, 2018.

Authors Profile

Dr. Mohammed mahmood ali, has completed his M.Tech in Soft Engineering from Jawaharlal Nehru Technological University, Hyderabad and Ph.D in Computer Science and Engineering from Osmania University respectively in 2007 and 2017 respectively. Currently working as Assosiate Professor in Department of Computer Science and Engineering, affiliated to Osmania University, Hyderabad. He is a member of IEEE & IEEE computer society since 2014 and life member of the IEI (India) since 2004. He has published more than 23 research papers in reputed International Journals indexed in Scopus, Thomson Reuters (SCI & Web of Science) and conference papers are indexed in IEEE and ACM. His main research work focuses on DATA Mining / Surveillance of Cyber messages, Mining Stress from Social media, Information analysis and Knowledge management. He has 10 years of teaching experience and 4 years of Research Experience.

Ms. Shaikha Hajera pursed B.Tech in Computer Science & Engineering from SWCET affiliated to Jawaharlal Nehru Technological University, Hyderabad in 2016. She is currently pursuing M.Tech in Computer Science and Engineering from MJCET affiliated to Osmania University. She is a member of Indian Society for Technical Education, since 2015. She has published two research papers, one in reputed international journal and other in conference including IEEE. Her main research work focuses on Natural Language Processing, Machine Learning, Data Mining, IoT and Affective Computing.